THIS PAGE HAS BEEN LEFT INTENTIONALLY BLANK

Chapter

10

Submission Tools

CHAPTER 10 – SUBMISSION TOOLS

The **Submission Tools** are comprised of the **Data Validation and Upload Tools**, which assist researchers with the validation and upload of data into the repository. The validation component verifies that submitted data conforms to the required format and range values defined in the Data Dictionary. **Note:** In certain BRICS instances, extra-validation rules are applied (See Appendix). The tool references the Data Dictionary and validates the data associated with the files identified by the user for submission into the Data Repository. It provides a report of any data discrepancies in the form of errors and warnings. The errors highlight when data does not conform to how that field has been defined. If errors are found, a submission package cannot be created. After successful validation, a submission package is created which allows the data to be submitted to the system with the upload component. The tool, which runs as a Java Web Start application, runs locally on a user's computer, requiring the Java runtime environment to be installed.

To ensure the quality of uploaded data and also to make data easy to query, data should be submitted in a specific format and range values should comply with the values defined in the data dictionary. It's critical to understand the specific format which is defined by the data elements and form structure you are submitting your data against. All submitted research data must be validated against the values defined in the data dictionary prior to submission. The form structure templates serve as the template for data submission. To facilitate this process, we provide the **Data Validation Tool** that assists researchers with the submission of their data.

The tool accepts CSV files from a researcher and validates the files' content against the values defined in the data dictionary. These CSVs must be in the form of the form structure templates available within the Data Dictionary. CSVs of data exported from an external system, such as redcap, that have not been formatted to match the form structure templates will not be recognized by the validation tool. For those CSV files that pass validation, the Data Validation Tool creates a submission ticket and submission package, both in XML format. After that, data are ready for uploading. The submission ticket is used by the **Upload Tool** to upload the data (in the form of a corresponding submission package) to the repository.

10.1 System Requirements

The **Data Validation Tool** and **Upload Tool** are available within the **Data Repository** module as part of the **Submission Tools** java web launcher. These tools run as a single application locally on your machine. **Note:** the most recent version of Java Runtime Environment (JRE) (7 or higher) is required in order to run the module. Make sure your computer has it installed.

10.1.1 Tool Input and Output

Tool Input:

- CSV files with clinical data or imaging metadata that conform to the form structure templates from the data dictionary.

Tool Output:

- A submission package and submission ticket (XML) ready for submission by the Data Upload Tool.
- An error log with validation errors and warnings (if any).

10.1.2 CSV Checklist

Before populating the form structure template in the form of a CSV file with data, check the downloaded template for the following:

- Ensure there are no edits to the template and that the form hasn't changed since it was downloaded.
- All **data elements** needed to capture data are included into the CSV file (and the form structure).
- The **form structure** short name is entered in the cell A1 in the CSV file.
- The **Main group** (with all data elements) is present in the CSV file.
- The group (if any) and data element names are the same as in the FS.

After verifying the above, ensure the following before submission:

- The column A, starting from the A2 cell has no data, but has marks (x) that represent the beginning of a new record.
- The columns that correspond to required DEs in the FS are populated with data (e.g., Main.GUID).

10.1.3 Submission Package

After the successful validation, the generated submission package includes:

- A submission ticket (XML)
- A data file (XML).

10.1.4 Validation Warnings

- If any validation errors or warnings are found, the tool provides a detailed report of any data discrepancies, errors, and warnings received.

- Validation warnings are just warnings and they do not prevent the creation of the submission package. If any validation errors are found, however, a submission package cannot be created. In that case, you should first edit the data to fix all errors, and then re-validate the data.

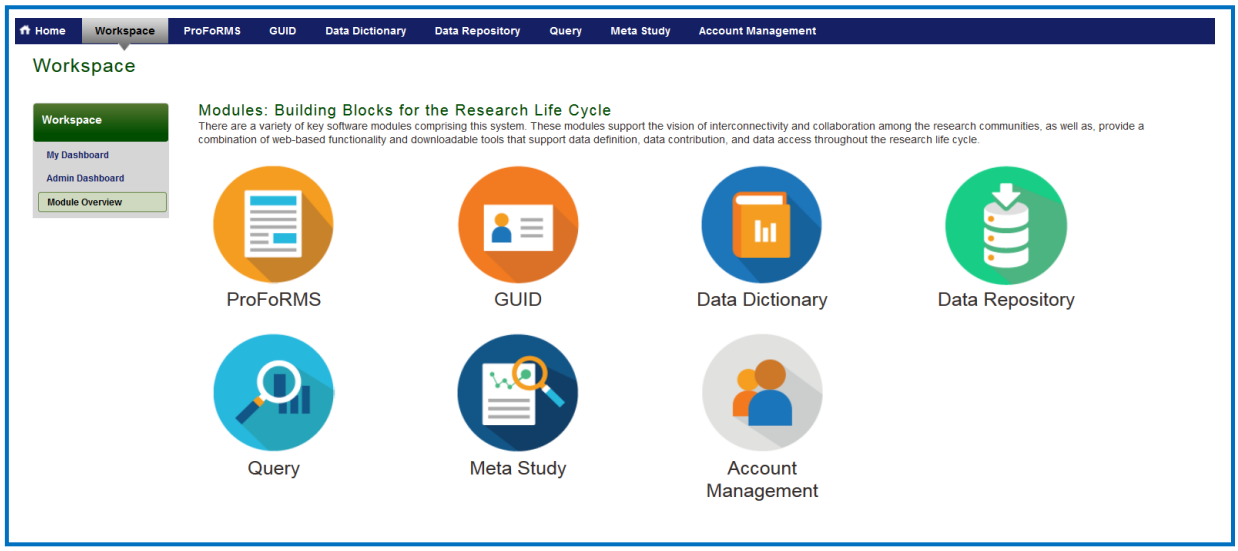
10.2 Data Validation Tool

The **Data Validation Tool** is available within the **Data Repository** module as part of the **Submission Tools**. For steps downloading the CSV template, refer to the **Data Dictionary** module.

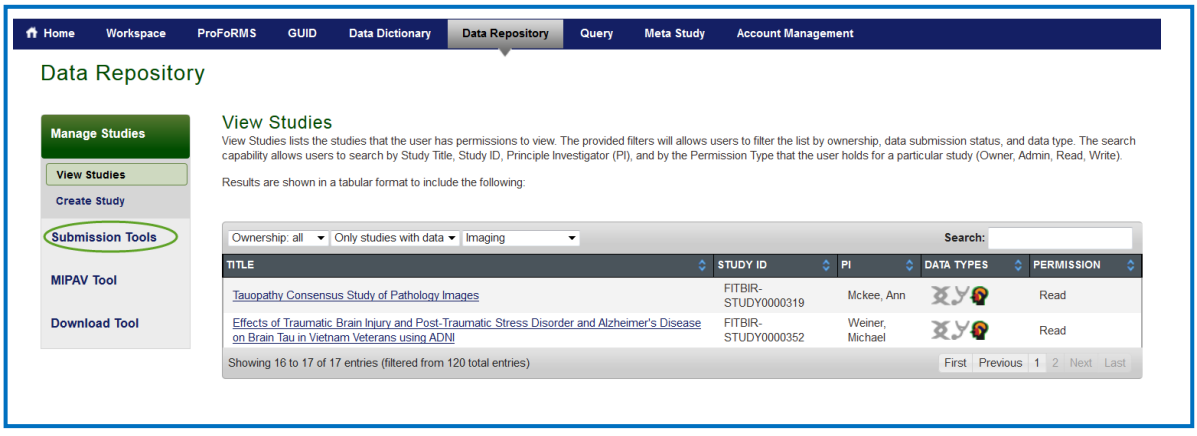
10.2.1 Running the Data Validation Tool

The Data Validation Tool runs locally on your machine. In order to launch the module,

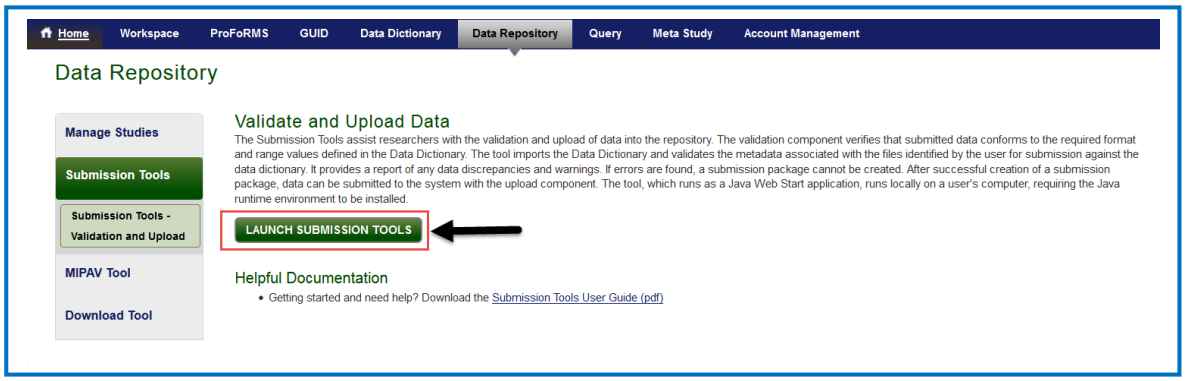
1. Log into BRICS, and navigate to your Workspace where all modules you have access to are displayed. **Note**: If the module is greyed out then you do not have access to the module. Please reach out to BRICS Operations to inquire about access should you not have access.



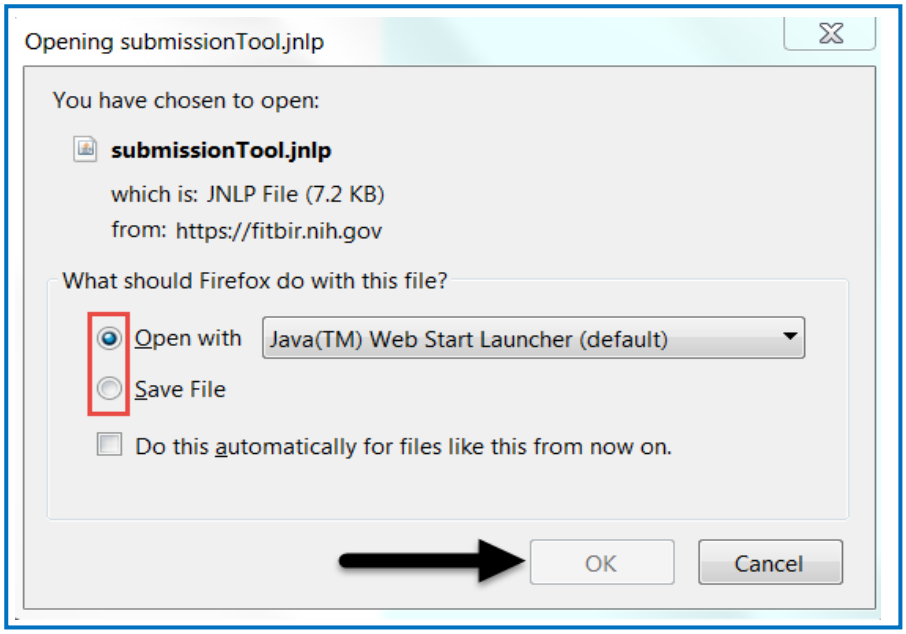
2. Click the **Data Repository** module from your Workspace.
3. Click the **Submission Tools** on the left-side menu.



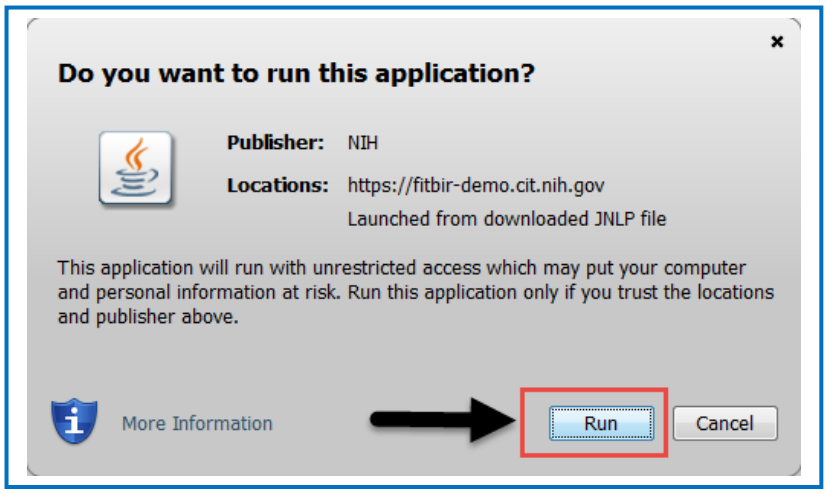
4. The **Validate** and **Upload Data** window appears. Click the **Launch Submission Tools** button.



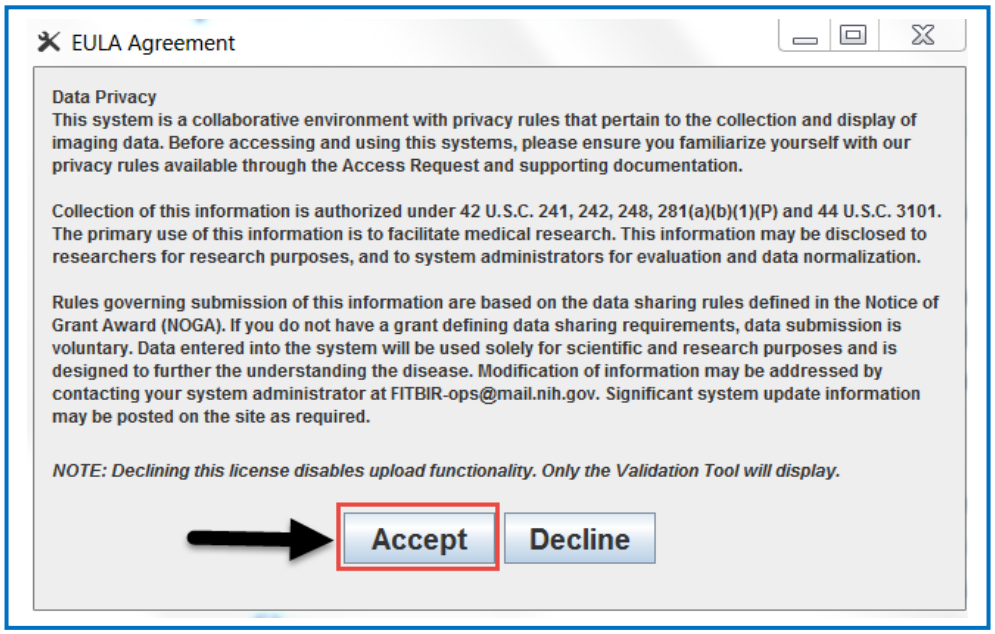
5. In the Opening window that appears, select **Open with Java(TM) Web Start Launcher (default)** and click the **OK** button.



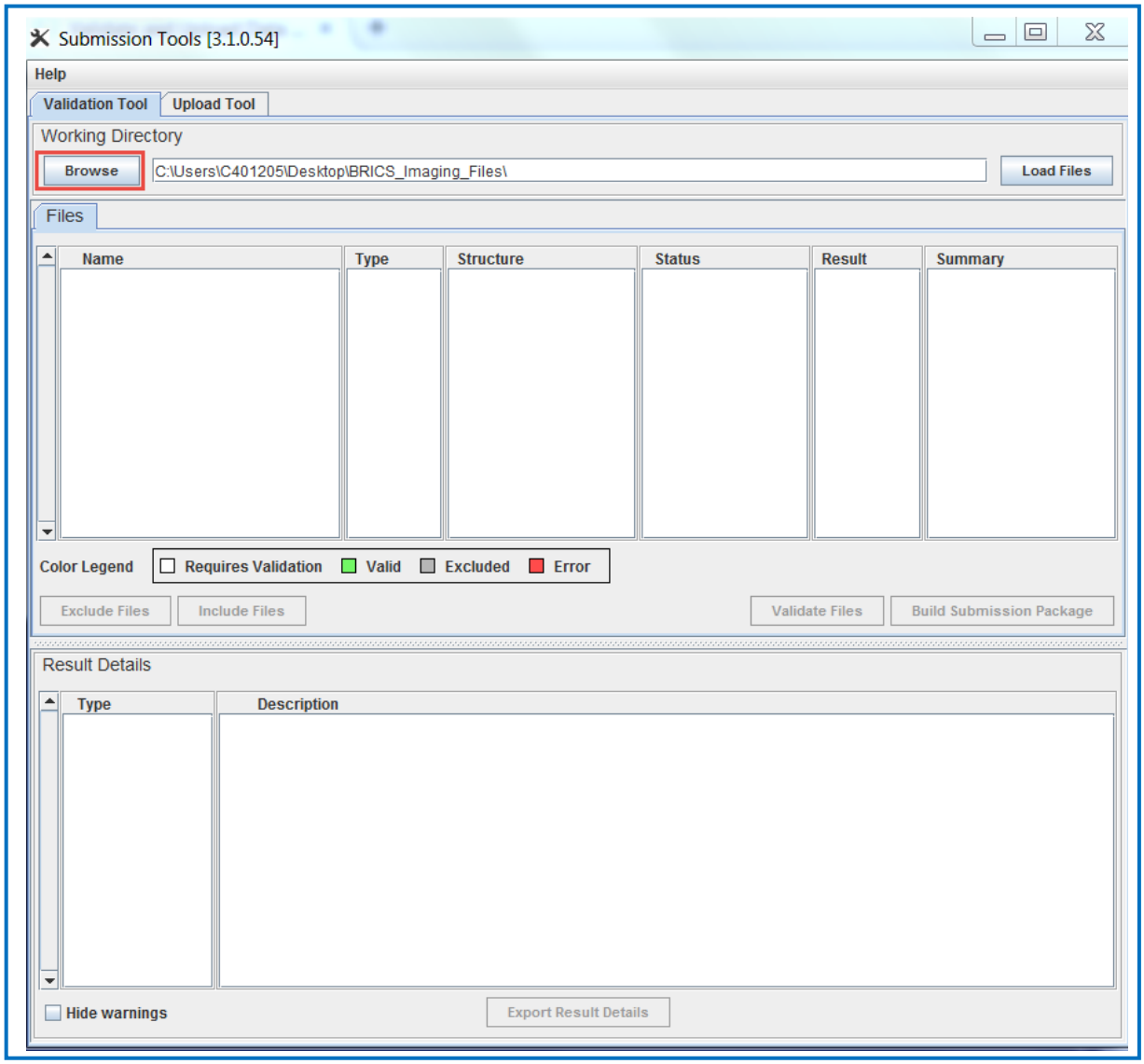
6. In the Java Runtime Environment window that appears next saying "**Do you want to run this application?**" click **Run**.



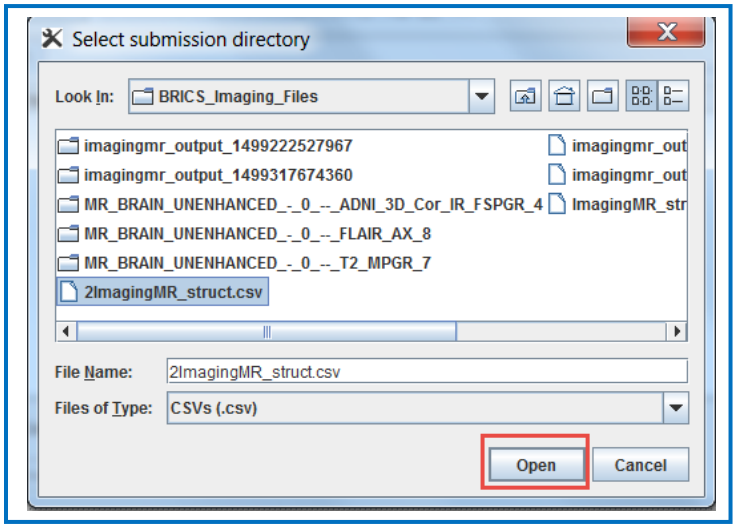
7. To continue, you **MUST** Accept the **EULA** agreement (this enables both the Validation and Upload Tools). If you click **Decline,** only the Data Validation Tool will be enabled.



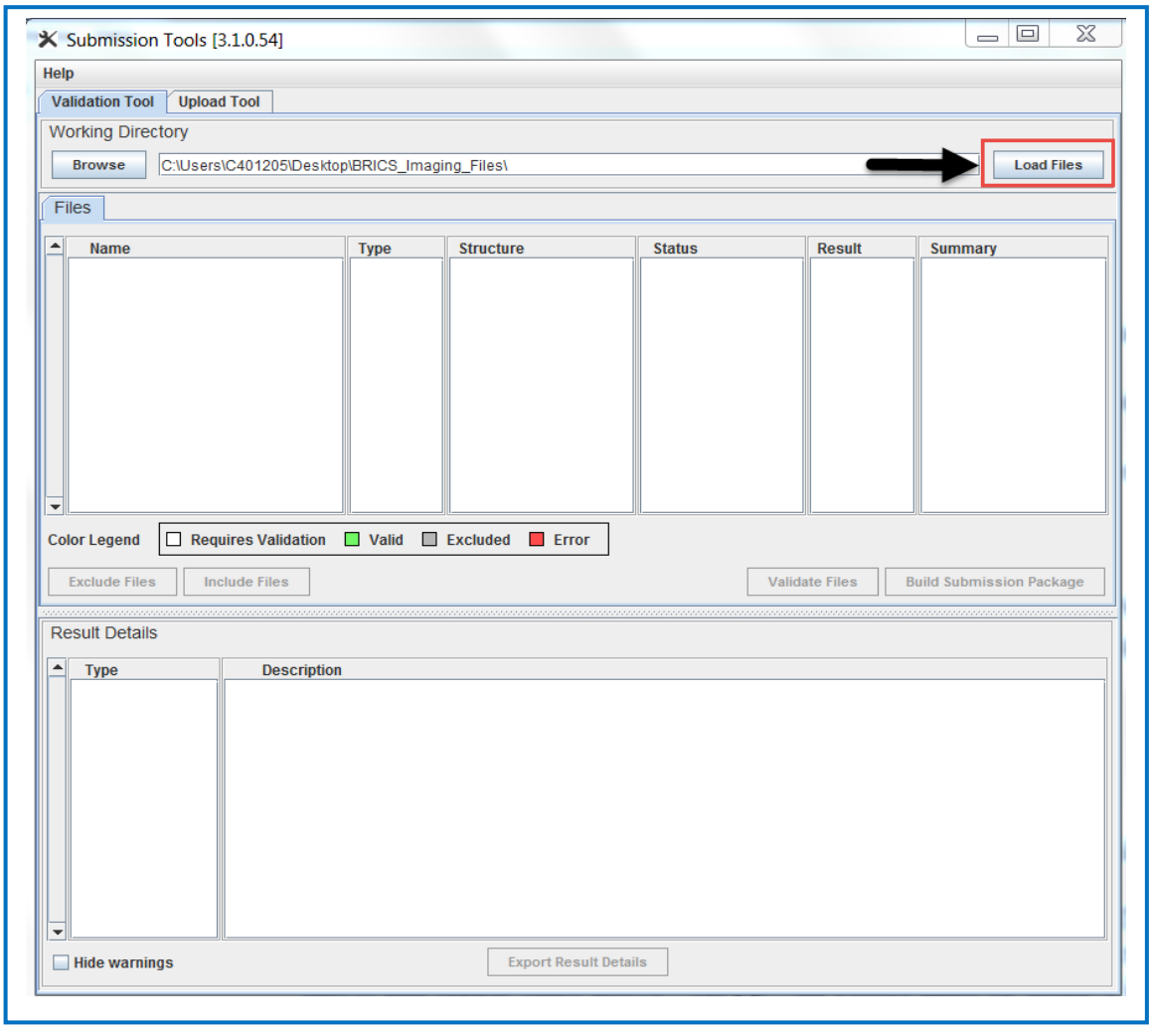
8. The Submission Tool appears. Click **Browse** (under Working Directory) to navigate to the directory where the files for submission (CSVs) are located. We call it your Working Directory. **Note**: It is recommended that you store CSV files in easy to access directories with only the necessary files for validation for fastest speed.



9. Select the directory and click **Open** to load CSVs into the dialog box. The Loading Files window appears showing the progress.



10. Click **Load Files**. All files in the directory will be displayed.



Excluding Files From Validation

To exclude files from validation, select individual file(s) (click to highlight) that are of TYPE UNKNOWN and those not needed for the submission. Hold Ctrl while clicking in order to highlight multiple files. Click Exclude Files. **Note**: Imaging files (ex. DICOM) should not be excluded in import.

Including Files For Validation

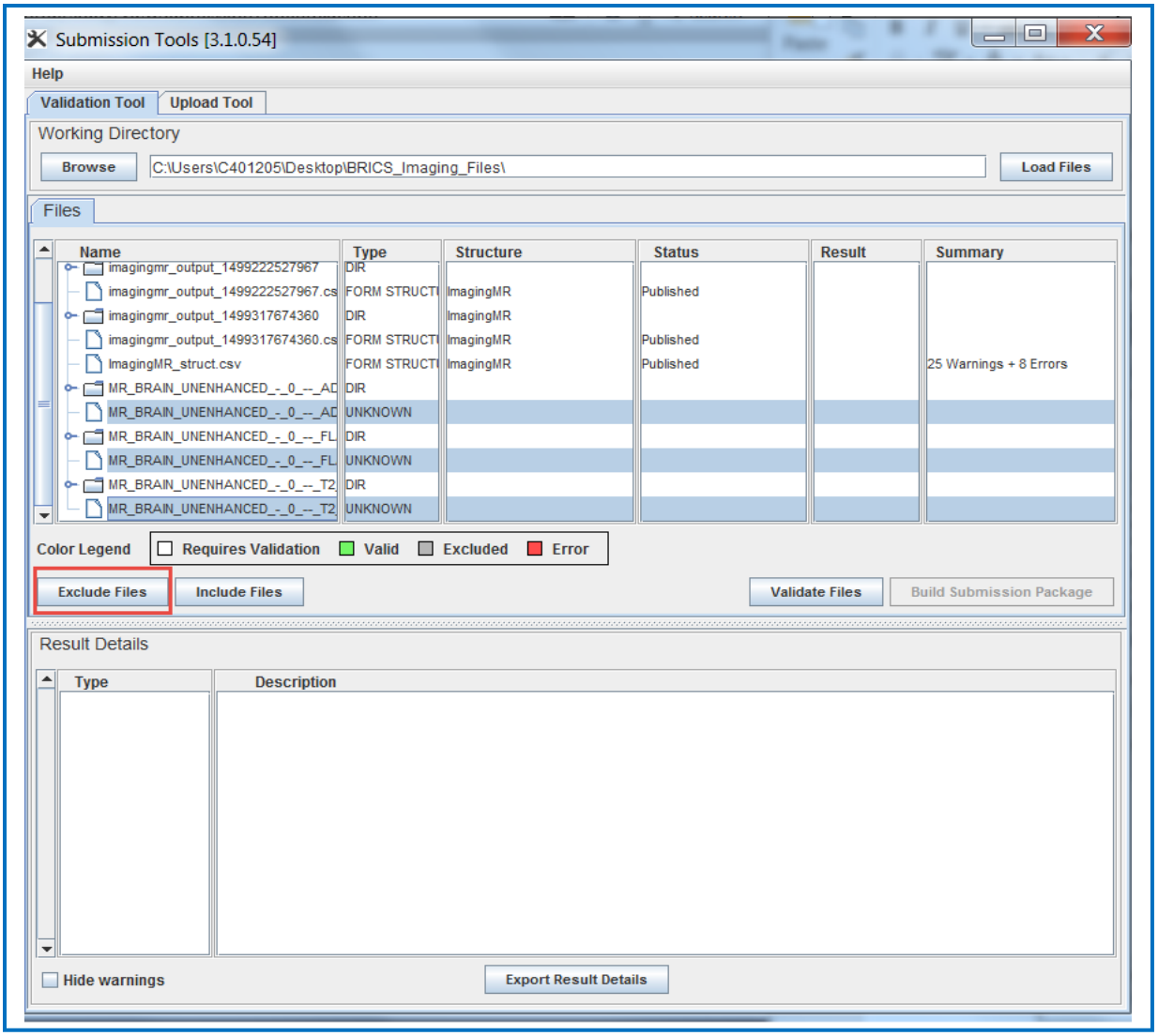
To include files for validation, select the CSV files you want to be validated and press Include Files. Hold Ctrl while clicking in order to highlight multiple files.

11. Select individual file(s) (click to highlight) that are of **TYPE UNKNOWN** and those files not needed for the submission. Hold Ctrl while clicking in order to highlight multiple files. Click **Exclude Files**. **Note**: Excluding files are not specific to unknown file types. There are instances where you are selecting 1 form to validate and submit and there are instances where the forms are saved incorrectly and show up as "Unknown".

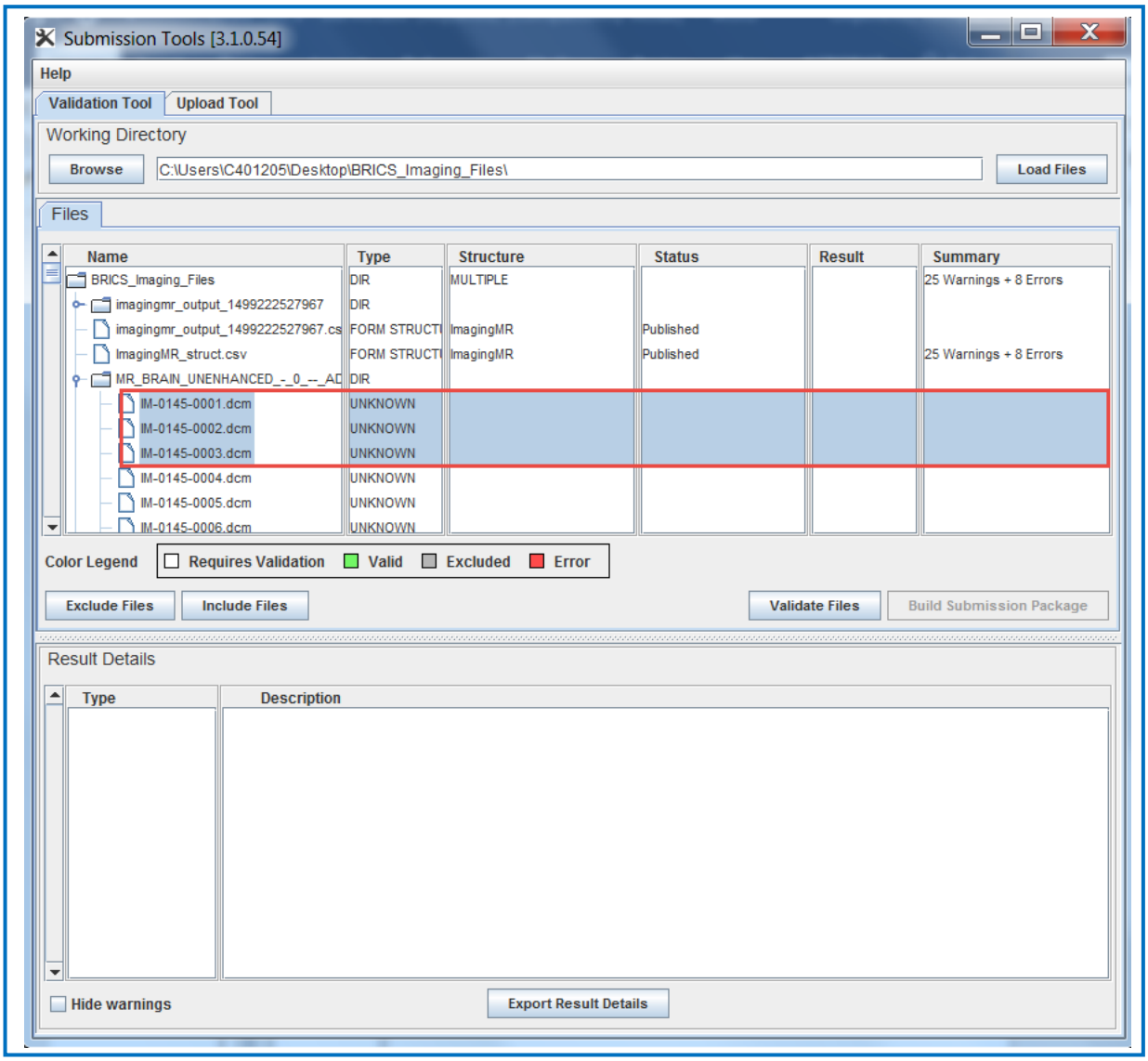
In an ideal world, your Working Directory should contain only CSV files for validation. Although very often it contains other files also (such as error logs and notes, etc.) In that case, you need to:

Exclude from validation those files (and directories) that are not designated for validation. These files usually appear with Type=UNKNOWN under Files in the Working Directory;

Include into validation the CSV data files that you would like to validate. These files have Type=CSV in the Files table.

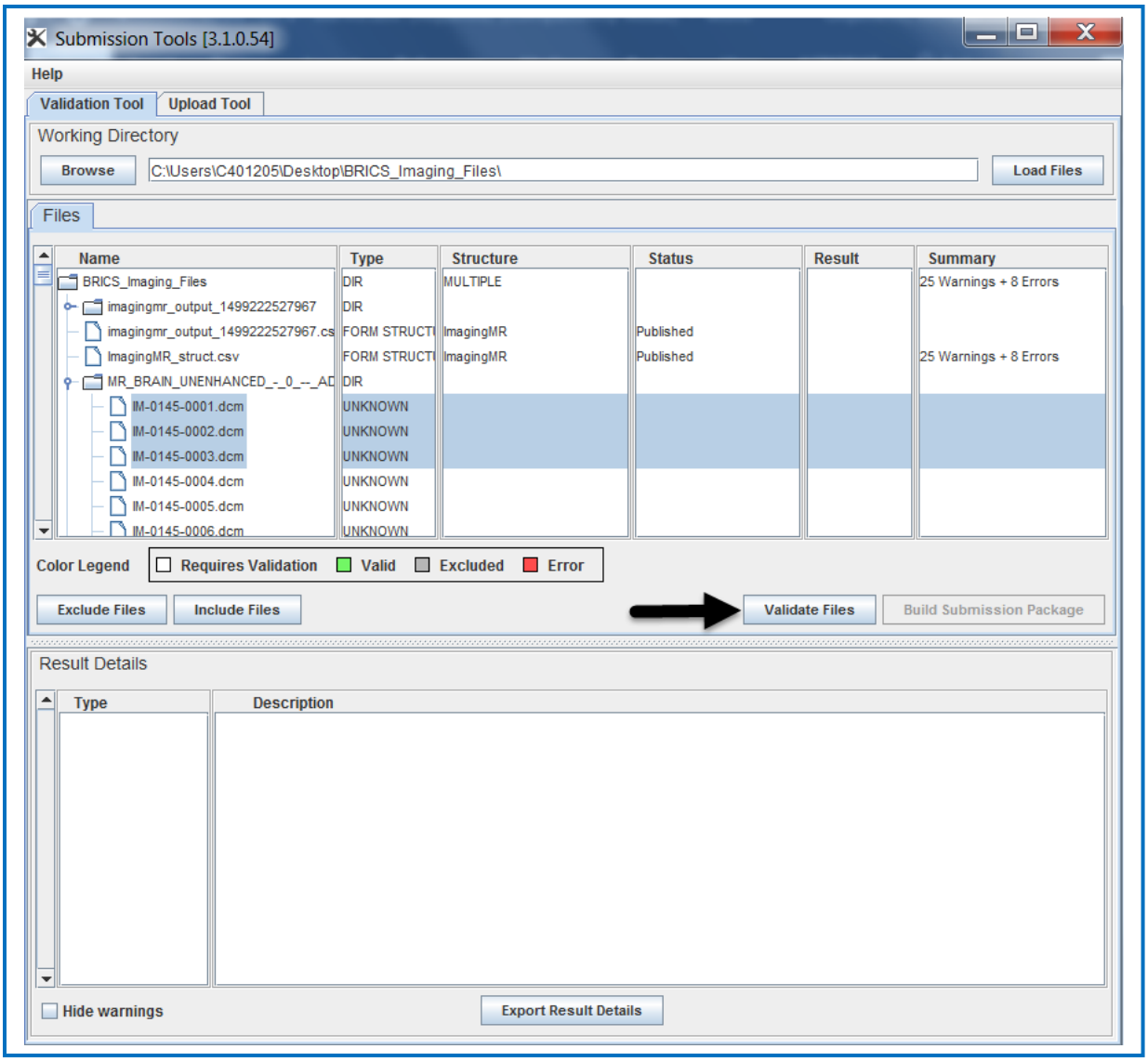

Submission Tools [3.1.0.54]
Help
Validation Tool
Upload Tool
Working Directory
Browse
C:\Users\C401205\Desktop\BRICS_Imaging_Files\
Load Files
Files
Name
Type
Structure
Status
Result
Summary
imagingmr_output_1499222527967
DIR
imagingmr_output_1499222527967.cs
FORM STRUCTI
ImagingMR
Published
imagingmr_output_1499317674360
DIR
ImagingMR
imagingmr_output_1499317674360.cs
FORM STRUCTI
ImagingMR
Published
ImagingMR_struct.csv
FORM STRUCTI
ImagingMR
Published
25 Warnings + 8 Errors
MR_BRAIN_UNENHANCED_-_0_--_AD
DIR
MR_BRAIN_UNENHANCED_-_0_--_AD
UNKNOWN
MR_BRAIN_UNENHANCED_-_0_--_FL
DIR
MR_BRAIN_UNENHANCED_-_0_--_FL
UNKNOWN
MR_BRAIN_UNENHANCED_-_0_--_T2
DIR
MR_BRAIN_UNENHANCED_-_0_--_T2
UNKNOWN
Color Legend
Requires Validation
Valid
Excluded
Error
Exclude Files
Include Files
Validate Files
Build Submission Package
Result Details
Type
Description
Hide warnings
Export Result Details

12. Select the remaining file(s) (click to highlight) to be validated. **All files that should not be validated should be in gray. Note:** In this example, 3 of the .dcm(dicom) files below are related to subjects' tests that are not meant to be submitted.



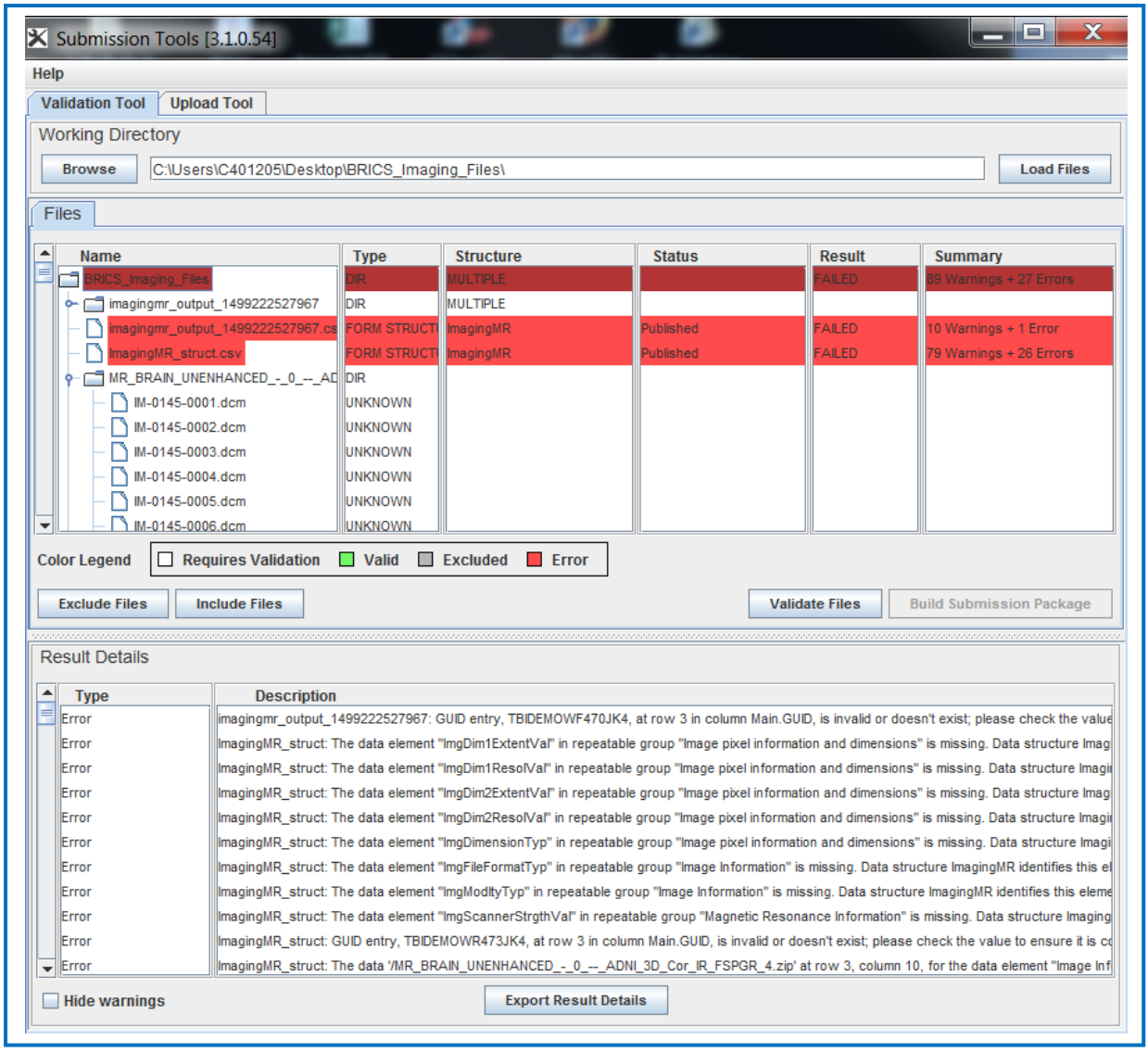
13. Click **Validate Files** button. If there are no errors, Click **OK** (All files are valid). The file(s) will then be highlighted in **GREEN** as Valid. If there are errors, Click **OK** (The validated files contain errors.) The file(s) will then be highlighted in **RED** as having an Error. Please make corrections to the file(s) and repeat steps above. **Note**: All Errors must be fixed. Files can pass validation with Warnings.

    Note: If there are no errors, you will see a prompt that reads "All files are Valid". Click **OK**." Same for if there are errors.



14. **Validation Results - A**: If there are errors found in your CSV file(s), for each file that fails validation: 1) the form structure name appears in the Structure column, 2) the word FAILED appears in the Result column, and 3) the summary column contains errors and possibly warnings. **Note**: a file that passed validation still can have a lot of warnings.

15. **Validation Results - B**: If there are no errors found in your CSV file(s), for each file that passes validation: 1) the form structure name appears in the Structure column, 2) the word PASSED appears in the Result column, and 3) the summary column may contain warnings, but no errors. **Note**: a file that passed validation still can have a lot of warnings.



10.2.2 Error Logs

Validation errors and warnings appear in the **Result Details** table. Files with warnings can be validated. Files with errors must be fixed and re-validated, and then resubmitted for another validation round.

Validation errors appear when a CSV file has entries that are:

- In different format (other than defined in the data dictionary for this data element),
- Not listed among permissible values for this particular data element,
- Left empty, but defined as required in the form structure,
- Have more than 1 permissible value separated by a semicolon ";" but the data element is not set to be a multiset.
- Do not pass extra-validation rules (See Appendix for a list of forms with extra-validation rules)

Note: Validation errors mostly appear when a data entry, which was defined as required in the corresponding form structure, is missing in the CSV file.

10.2.3 Fixing Validation Errors

Validation errors and warnings appear in the Result Details table. Files with warnings can be validated. However, files with errors must be fixed and re-validated, and then resubmitted for another validation round.

Validation errors and warnings can be exported into a text file - that makes working with them and fixing errors much easier. To export validation errors or warnings, or both:

1. Click the **Export Result Details**.
2. In the Save dialog box that appears:
    a) select a directory where you would like to save validation logs,
    b) specify what types of error log entries you would like to export. These could be a) both errors and warnings (recommended only for smaller log files),
    c) errors only (recommended), or c) warnings only.
3. Type in your own file name and Click **Save**.
4. The log file will be saved in the designated directory under the chosen name.

10.2.4 Best Practices

By default, an error log file is created and stored in the same directory as your working files. We recommend that you create a designated error log directory and save validation logs there.

By default, an error log is saved under the "resultDetail.txt" name. We commend that you choose your own file name for an error log and that the name is somehow related to the name of your data file. E.g. if you have a data file, let's say "MyData.csv", you give the corresponding error log file the following name "MyDataErrorLog.txt".

After you have exported all validation errors:

1. Open the log file in a text editor (MS Word, Notepad, Crimson, Notepad++ - all these will work).
2. Open your CSV file in MS Excel or your preferable text editor that can work with CSV (not MS Word!).
3. Go through each entry in the error log and fix it in the CSV file. Save changes in the CSV file. Make sure you saved it as CSV. Some of these changes may require additional review and consultation with your principle investigator. If you have an error for data that you cannot change, please reach out to your BRICS point of contact.
4. Re-validate the fixed CSV file. Make sure that all errors are gone.
5. Create the submission package.

If you received numerous validation errors, we recommend that you work on fixing them in batches. Fix a few errors, save the fixed CSV file and re-run it through the Data Validation Tool. It will still give you a lot of errors, but we hope it would be fewer that before. Save the new error log and go through it fixing a few more errors. Re-run validation. Repeat these steps until you get 0 (zero) errors.

10.3 DATA UPLOAD TOOL

The **Data Upload Tool** is available within the **Data Repository** module as part of the **Submission Tools**. The Data Upload Tool helps researchers to upload data to the data repository. This phase is not accessible without a dataset that passed validation. The data should be uploaded in the form of a submission package (XML) that has a unique identifier - a submission ticket (XML). If you don't have a submission package ready, use the Data Validation Tool to create it. The tool will also validate your data and make sure that they conform to the required format and range values defined in the data dictionary.

10.3.1 Running the Data Upload Tool

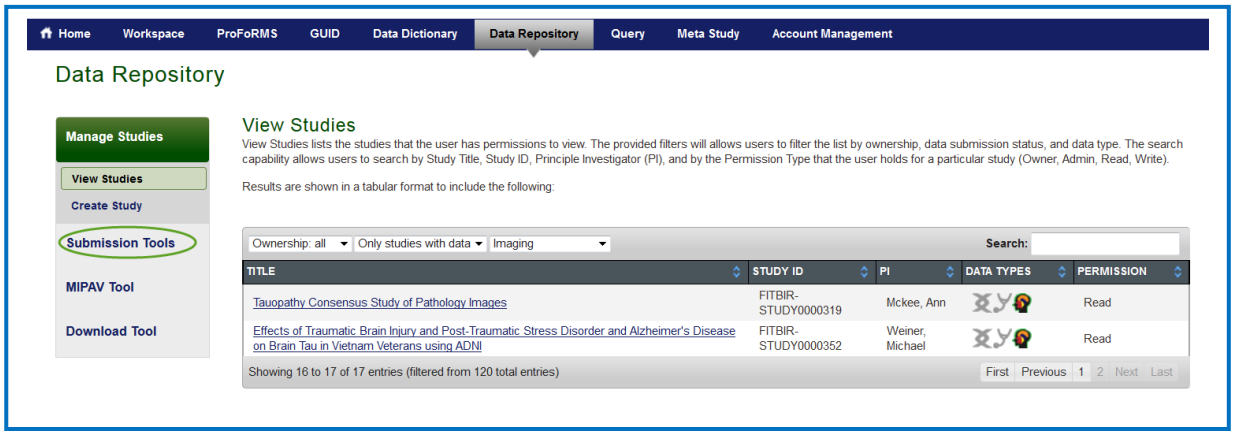
The Data Upload Tool runs locally on your machine. In order to launch the module,

Perform the following actions:

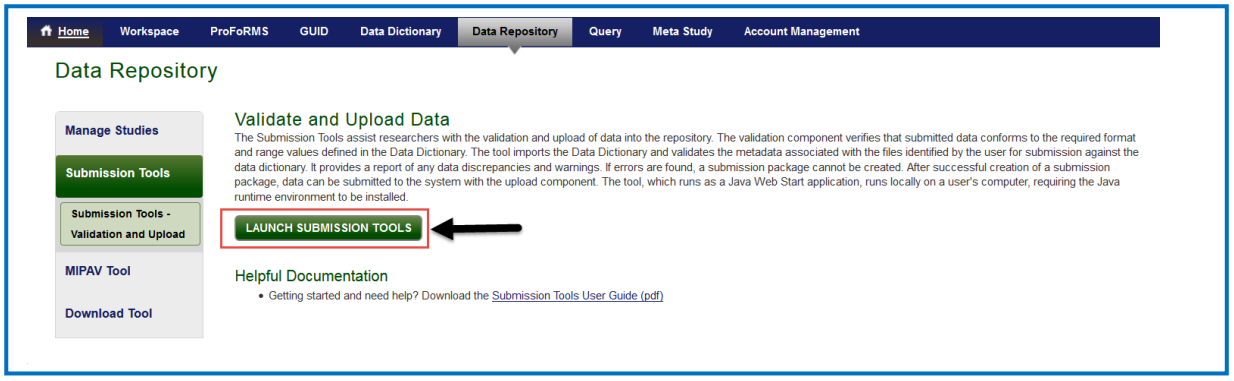
1. Navigate to your Workspace,

2. Click the **Data Repository** module from your Workspace,

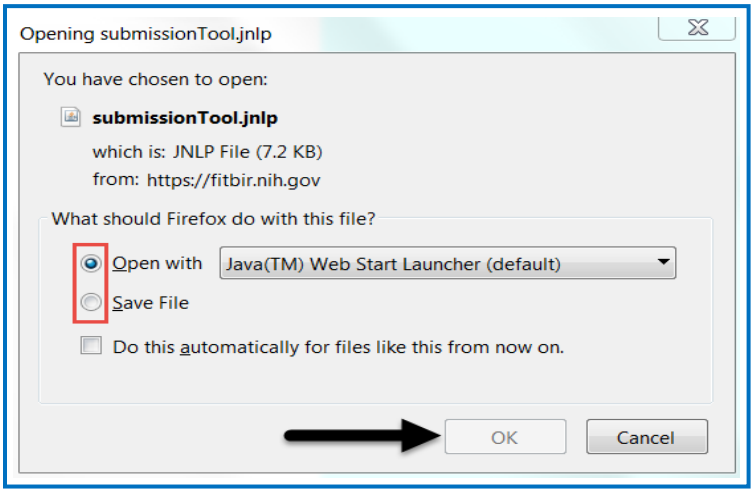
3. Click the **Submission Tools** on the left-side menu.



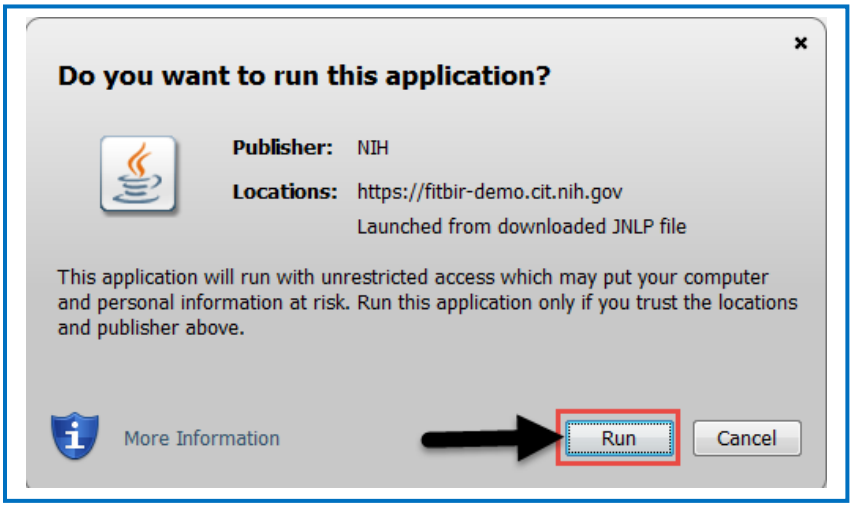
4. The **Validate** and **Upload Data** window appears. Click the **Launch Submission Tool** button. The Java Web Start Launcher opens



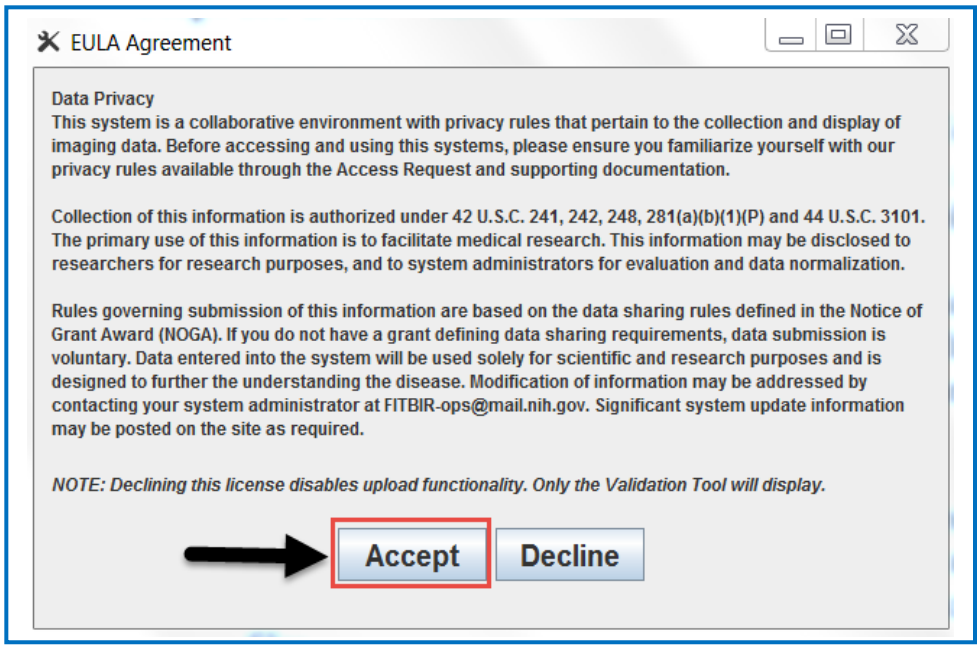
5. In the Opening window that appears, select Open with Java(TM) Web Start Launcher (default) and click **OK** to launch the Submission tool.



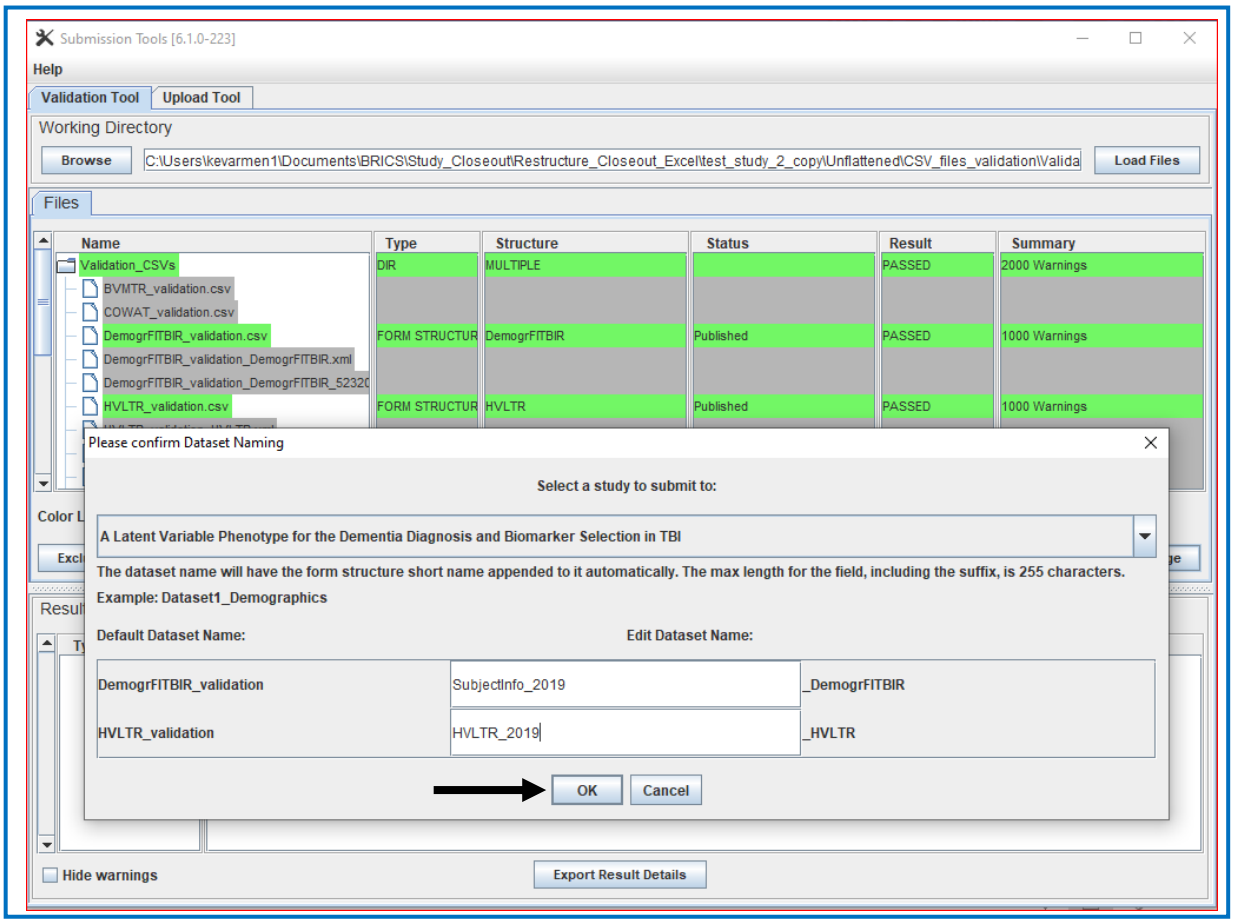
6. In the Java Runtime Environment window that appears next saying **"Do you want to run this application?"** Click the **Run** button.



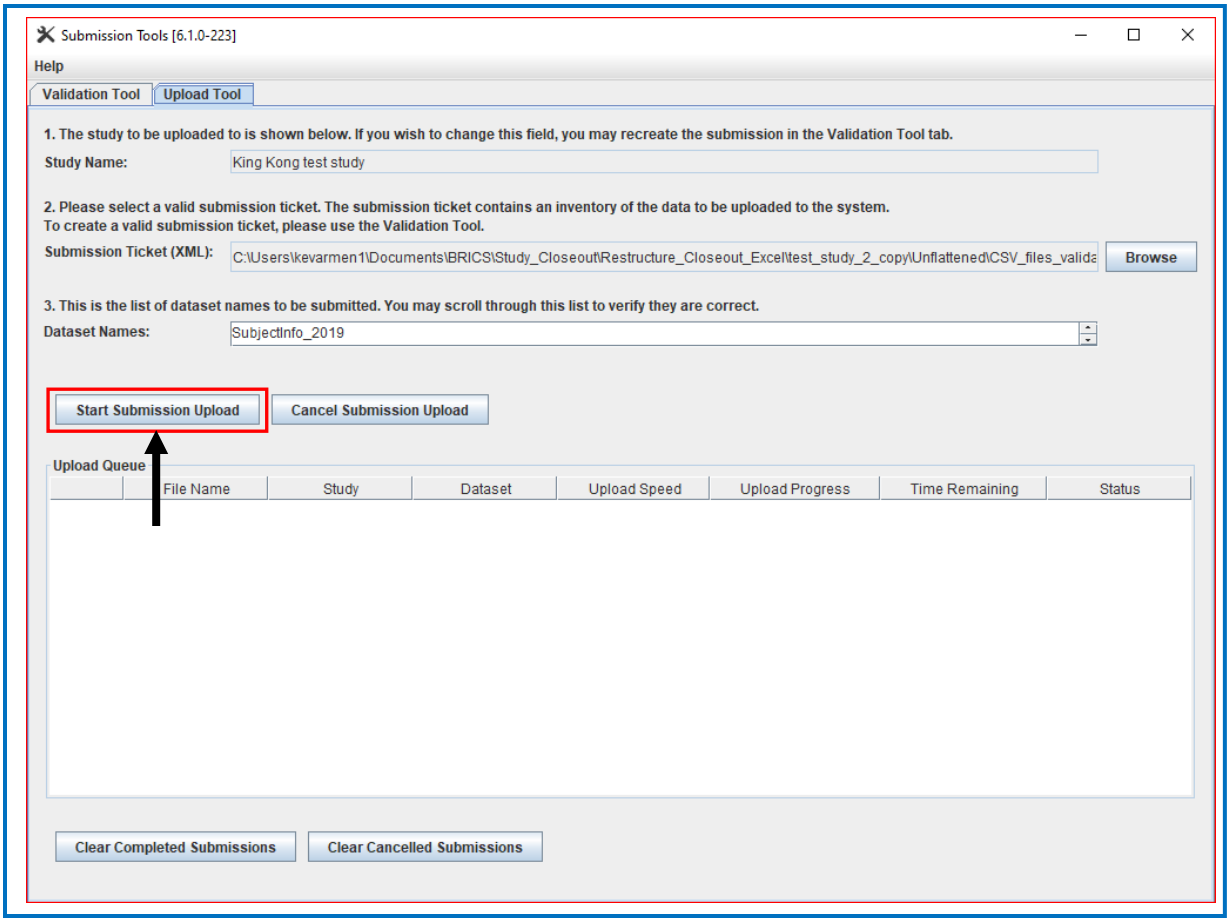
7. The EULA Agreement window appears displaying the data privacy user agreement. Read the agreement and click Accept if you agree (this enables both the Validation and Upload Tools). If you click the **Decline,** only the Validation Tool will be enabled.



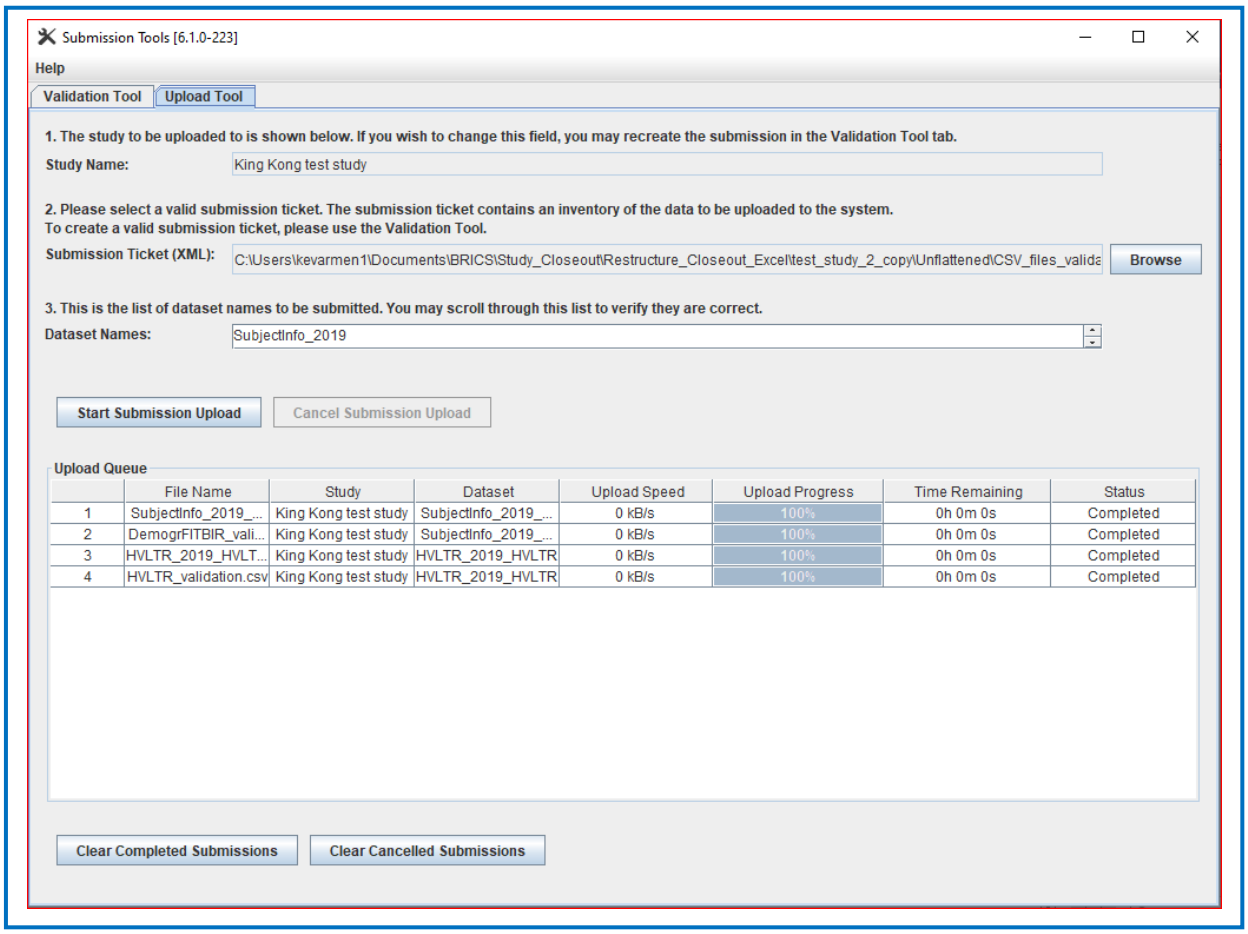
8. After files have been successfully validated, click **Build Submission Package**.
   a) Use the drop-down Study Name menu to select the study name.
   b) Enter dataset names in window. The form structure short name will be appended to the end of user assigned dataset name, e.g., SubjectInfo_2019_DemogrFITBIR, where "SubjectInfo_2019" is the user assigned dataset name and "DemogrFITBIR" is the form structure short name.
   c) Click **OK** button.



9. If accessing the **Upload Tool** immediately following the Validation Tool, **item 2** of the Upload Tool is auto-populated. Verify that the correct study was selected to upload the data to. Once information for Steps 1-3 have been confirmed, Click **Start Submission Upload. Note**: When you relaunch and attempt to go straight to "**Upload**" you will have to first **Re-validate** the data files.



10. The submission package appears in the **Upload Queue** table. In the Upload Queue table, you can watch the progress of your submission(s). The table will update as file(s) are being uploaded to the system. For successful upload(s), the Status (see the Status column) will be designated as "**Completed**".



ICON KEY

Notes

Important Information

Things to Note:

- If you need to cancel your submission, Click the Cancel button;
- To clear the completed submissions list, use the Clear Completed Submissions button located at the bottom of the Upload Manager window.
- To clear the cancelled submissions list, use the Clear Cancelled Submissions button located at the bottom of the Upload Manager window.

10.3.2 Best Practices

After submitting the data, to make sure that your dataset appears under the study you selected:

1. Navigate to the Data Repository -> View Studies page.
2. Find your study on the study table. **Note**: 3 icons located in the Data Types column. If your study has any data submitted, at least one of the icons appears in color.
3. Select the study and click on the study name to open the Study Overview page.
4. On the study page, click on the "+" sign next to Dataset Submissions.
5. The table that contains all submitted datasets opens. Make sure that your dataset is listed in this table.
6. Contact the BRICS System Administrator if you do not see the data you uploaded

The data types associated with a study are represented by three icons:

1. **Double helix** - represents genomics data;
2. **Stethoscope** - represents clinical assessment data;
3. **Head profile** - represents imaging data.

If the icons next to the study name are highlighted in color, the study has datasets of the highlighted types.

ICON KEY

Notes

Important Information

Things to Note:

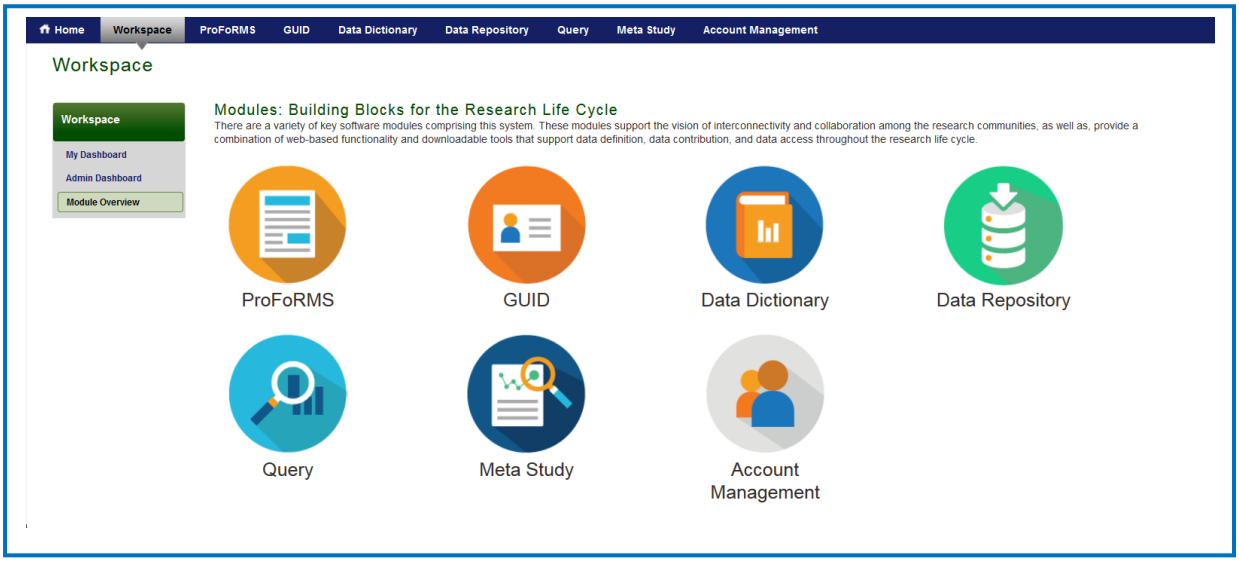
- You can only upload data to the studies you have the data upload permissions.
- To make sure that the most recent list of studies is available for you, use the Refresh button to update the list of studies.
- The name assigned to the uploading dataset must be unique for the selected study.

10.3.3 Adding Submitted Data to the Download Queue

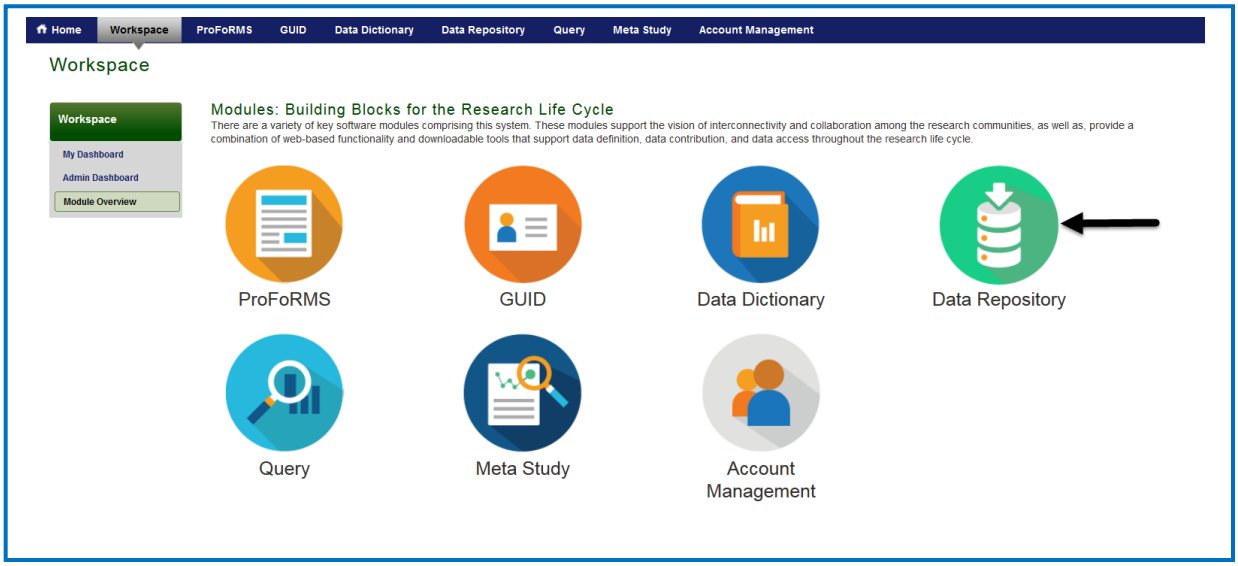
The Data Download feature helps users to select and download datasets from the data repository to their own computer systems. Users may select datasets to download by placing them in their Download Queue (see below). Then, using the **Download Tool**, users may download selected datasets from the repository to their own systems.

To Populate Download Queue: Perform the following actions:

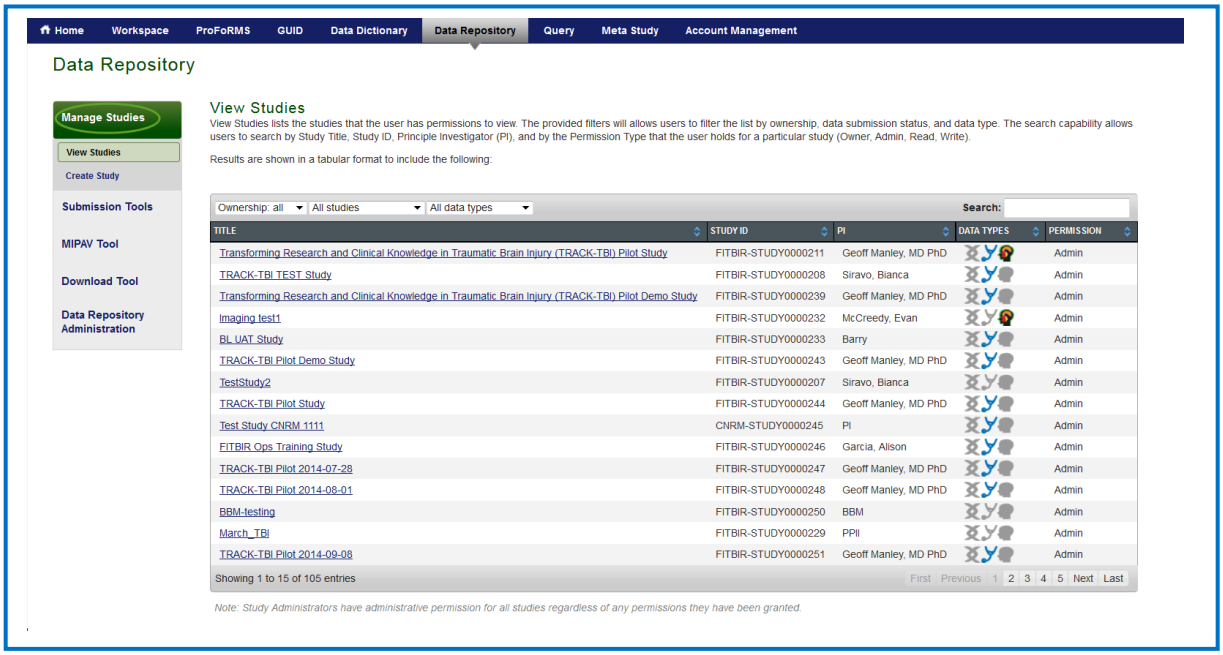
1. Login to the system.
2. Navigate to the Workspace landing page



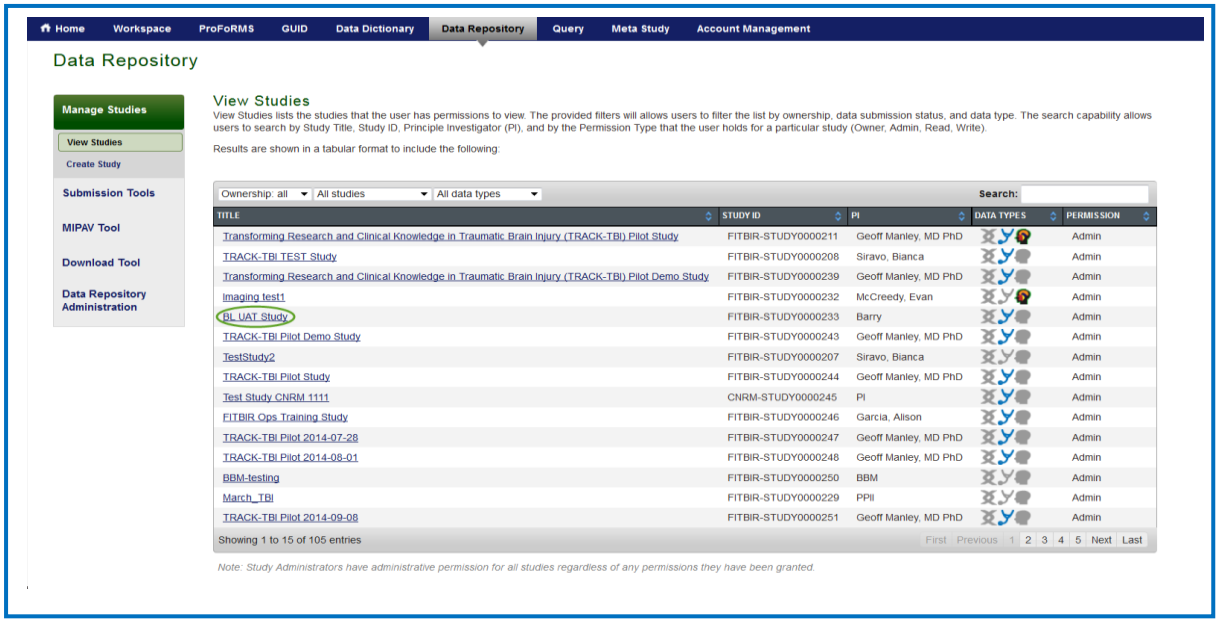
3. Click the **Data Repository** module.



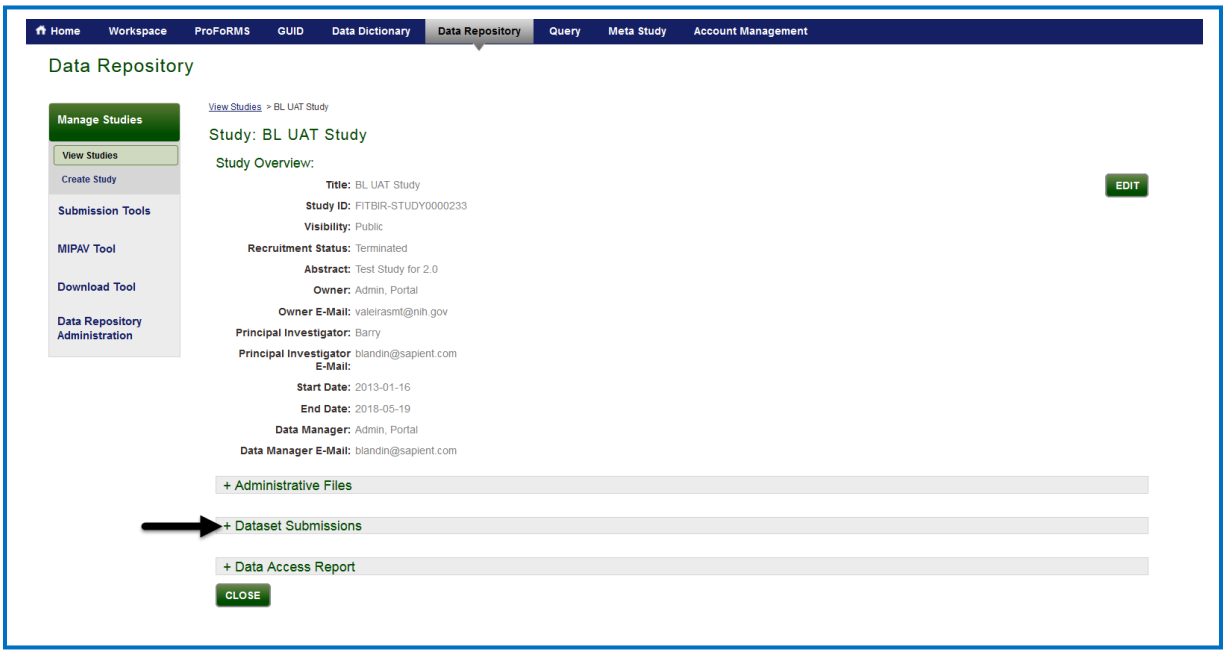
4. Click the **Manage Studies** on the left-side tool bar. The View Studies list opens with the menu options.



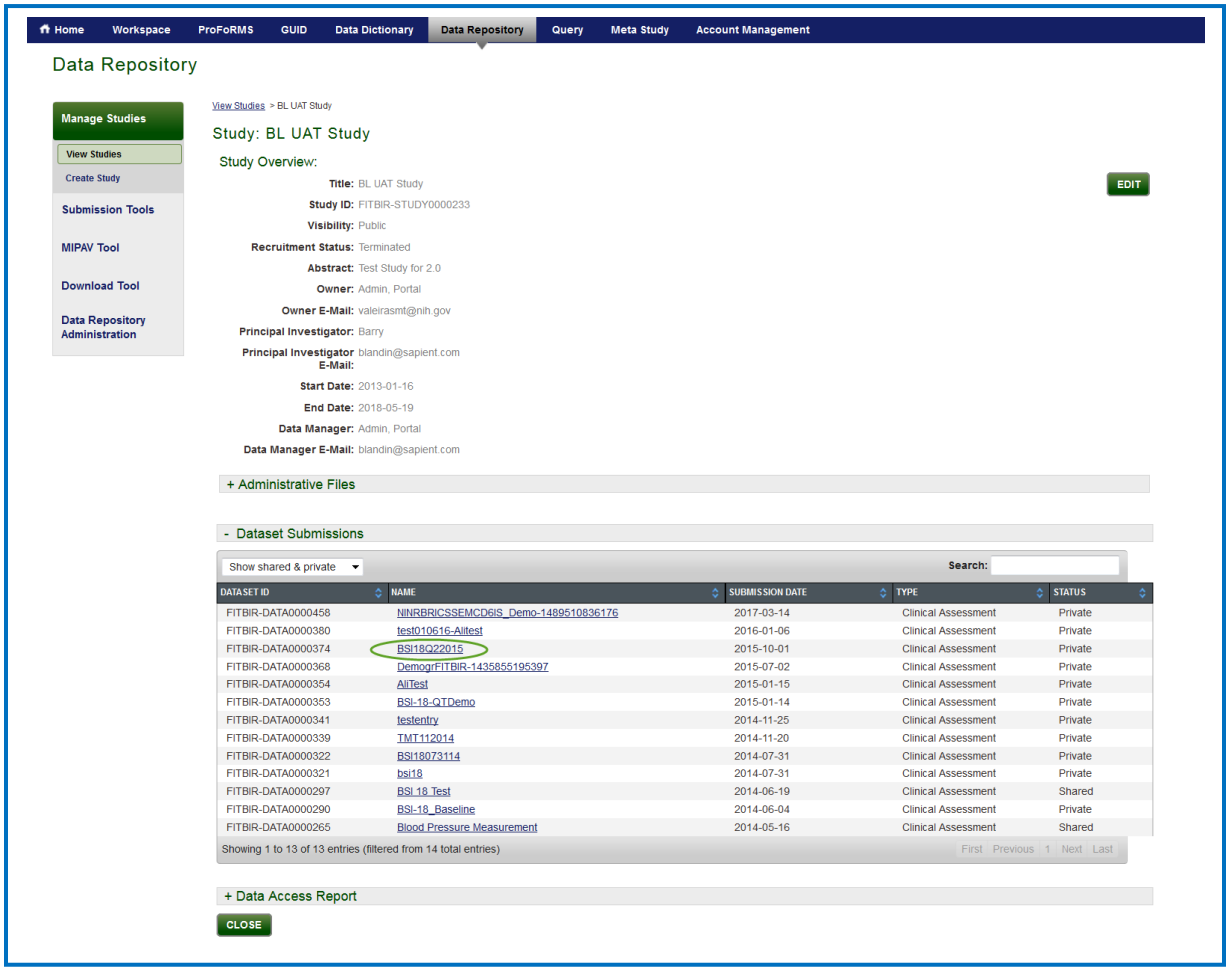
5. Select the **Study** from which you wish to download a dataset by clicking on the **Study Title**. **Note**: *View Studies lists the studies that the user has permissions to view.* The provided filters will allow users to filter the list by ownership, data submission status, and data type. The search capability allows users to search by: (***Study Title, Study ID, Principle Investigator*** *(PI), and by the* ***Permission Type***) that the user holds for a particular study (Owner, Admin, Read, and Write).



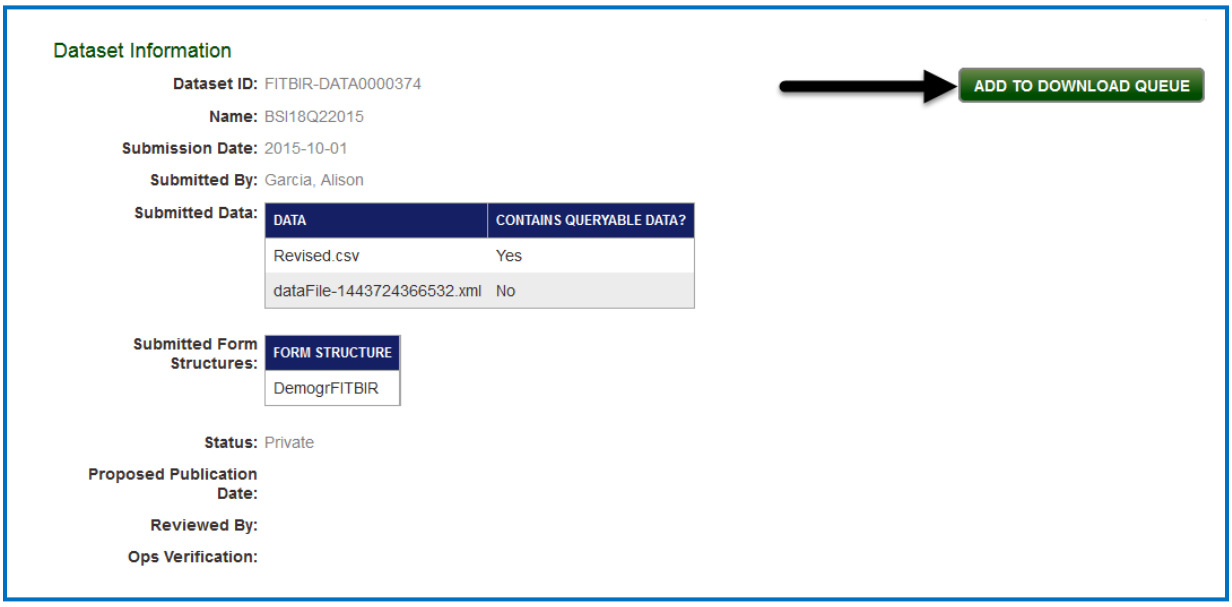
6. Click on the + next to **Data Set Submissions**.



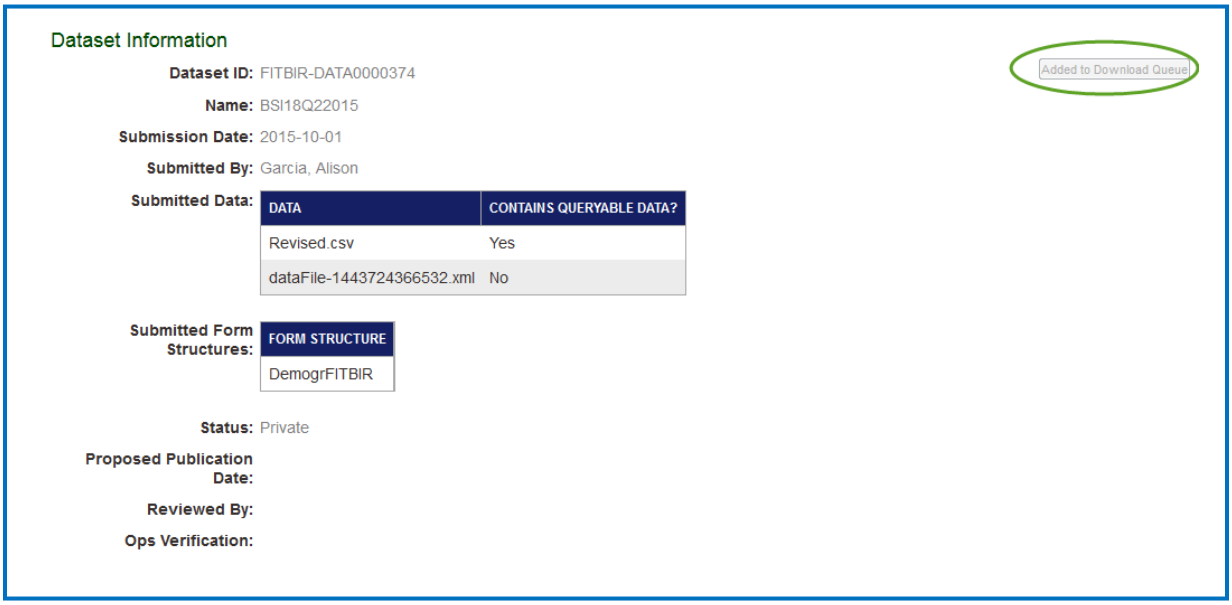
7. Select the **Dataset** you wish to download by clicking on its name. A **pop-up** window opens with information regarding the dataset.



8. Click **Add To Download Queue**.



9. The Dataset is **Added to Download Queue**. You will receive an e-mail indicating "***Requested data is now available in the Download Manager***".



APPENDIX: EXTRA-VALIDATION RULES

In select instances of BRICS (FITBIR), extra-validation rules have been incorporated into the **Validation Tool** for a subset of standard forms in the Data Dictionary. **Note**: this is in addition to the baseline validation checks that apply to all forms. Validation rules for additional standard forms will be added with future updates. The Extra Validation rules were developed based on the scoring guidelines and algorithms of the standard forms. The Extra Validation rules will provide an additional quality check for users when submitting data, therefore providing better quality of data for the TBI research community.

The Validation Tool will check for:

a) Submitted summary scores that have been miscalculated.
b) Submitted Summary scores with missing item level data.
c) Submitted completed item level data with missing summary scores.
d) Submitted age inputs not meeting the age requirement for the form

If any extra-validation errors are found, a submission package cannot be created. The error log will inform you of the specific extra-validation rules that are not met and where in the CSV they are located. You should edit the data to fix all errors, and then re-validate the data

Extra-validation rules were created uniquely for the form structures listed below:

1. **Alcohol, Smoking, and Substance Use Involvement Screening Test (ASSIST)**
2. **Alcohol Use Disorders Identification Test - Consumption Questions (AUDIT-C)**
3. **Alcohol Use Disorders Identification Test: Self-Report Version (AUDIT)**
4. **Balance Error Scoring System (BESS)**
5. **Beck Anxiety Inventory (BAI)**
6. **Beck Depression Inventory II (BDI-II)**
7. **Brief Symptoms Inventory-18 (BSI-18)**
8. **Brief Visuospatial Memory Test - Revised (BVMT-R)**
9. **Controlled Oral Word Association Test (COWAT)**
10. **Deployment Risk and Resilience Inventory (DRRI-2), V2 Combat**
11. **Disability Rating Scale (DRS)**
12. **Dizziness Handicap Inventory (DHI)**
13. **Drug Abuse Screening Test (DAST-10)**
14. **FIM Instrument**
15. **Glasgow Coma Scale and Pupils (GCS)**
16. **Glasgow Outcome Scale - Extended (GOS-E)**
17. **Grooved Pegboard Test (GPT)**
18. **Headache Impact Test (HIT-6)**
19. **Immediate Post-Concussion Assessment Testing (ImPACT)**
20. **Insomnia Severity Index (ISI)**
21. **King-Devick Concussion Screening Test**
22. **Mayo-Portland Adaptability Inventory-4 (MPAI-4)**
23. **Movement Disorder Society - Unified Parkinson's Disease Rating Scale (MDS-UPDRS)**
24. **Montreal Cognitive Assessment (MoCA)**

25. Neurobehavioral Symptom Inventory (NSI)
26. Patient Health Questionnaire - 9 (PHQ-9)
27. Pittsburgh Sleep Quality Index (PSQI)
28. Posttraumatic Stress Disorder Checklist - Civilian Version (PCL-C)
29. Rivermead Post-Concussion Symptoms Questionnaire (RPQ)
30. Satisfaction with Life Scale (SWLS)
31. 12-Item Short Form Health Survey Version 2 (SF-12v2)
32. 36-Item Short Form Health Survey (SF-36) version 2
33. Trail Making Test (TMT)
34. Vestibular/Ocular Motor Screening (VOMS)

.

THIS PAGE HAS BEEN LEFT INTENTIONALLY BLANK